Invisible Shadow for Navigation and Planning in Minimal Invasive Surgery

Marios Nicolaou, Adam James, Benny P.L. Lo,
Ara Darzi, and Guang-Zhong Yang

Royal Society/Wolfson Medical Image Computing Laboratory & Department of Surgical
Oncology and Technology, Imperial College London, London, United Kingdom
{m.nicolaou, a.james, a.darzi, g.z.yang}@imperial.ac.uk

Abstract. Depth estimation is one of the most fundamental challenges for performing minimally invasive surgical (MIS) procedures. The requirement of accurate 3D instrument navigation using limited visual depth cues makes such tasks even more difficult. With the constant expectation of improving safety for MIS, there is a growing requirement for overcoming such constraints during MIS. We present in this paper a method of improving the surgeon's perception of depth by introducing an "invisible shadow" in the operative field cast by an endoscopic instrument. Although, the shadow is invisible to human perception, it can be digitally detected, enhanced and re-displayed. Initial results from our study suggest that this method improves depth perception especially when the endoscopic instrument is in close proximity to the surface. Experiment results have shown that the method could potentially be used as an instrument navigation aid allowing accurate maneuvering of the instruments whilst minimizing tissue trauma.

1 Introduction

Over the last decade, minimally invasive surgery (MIS) has attained great popularity and acceptance among surgeons and patients alike due to its improved cosmetic appearance, shorter rehabilitation, less pain and decreased hospital costs. However, MIS requires a higher degree of competency from the surgeon due to the presence of a number of constraints. Among them, vision is the primary element. The surgeon is required to reconstruct the 3D operative field and perform instrument navigation through the narrow monoscopic two-dimensional (2D) field of view provided by the endoscope. The perceptual cues that a surgeon uses to navigate are complex, and it is well understood and documented in *cue theory* that a variety of cues are utilized in order to estimate depth. The visual system typically infers depth based on information relating to the posture of the eyes as well as visual patterns projected onto the retina [1]. The particular cues that foster the perception of depth have been widely investigated [2,3] and are often classified as *primary* (physiological) cues, such as binocular disparity, convergence, and accommodation, and *secondary* (pictorial) cues, such as linear perspective, elevation, shading and shadow, texture and texture gradients, and reference frames [1]. It is not well understood, however, how these cues are assimilated in MIS. This is because the majority of cues are subtle and difficult to detect in the operative environment presented to the surgeon. Furthermore,

there is a direct reduction in perceptual information that is available at any moment in time due to an ever-evolving surgical scene as a result of laparoscope translations and anatomical deformations. Ultimately, the monoscopic field-of-view provided by the laparoscope limits the 3D perception by presenting a scene onto 2D planes. It has been observed that surgeons tend to compensate for the lack of depth perception by developing new strategies such as groping forward and backward with instruments to gauge the relative depths of organs by touching them (Figure 1). The combined visual and haptic feedback helps to confirm the instrument position and orientation. This navigation approach, however, is not ideal particularly when undertaking delicate surgical maneuvers that require subtle control of instruments which must be performed slowly enough to avoid damaging the tissues in contact [4].
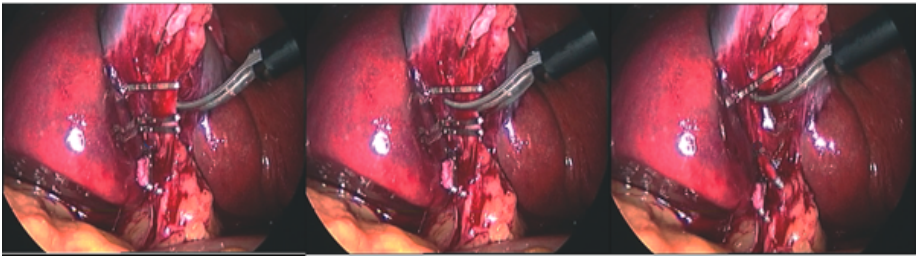

Fig. 1. Operative stills representing a collision sequence during MIS. In this sequence, the surgeon advances the instrument into the vessel to establish position before straddling the vessel instrument to transect it.

Currently, there is a constant requirement for surgery to become safer, particularly in the current climate of clinical governance. Practically, safety can be achieved by better training as well as by reducing the constraints set by the nature of MIS. Improving 3D visualization and ultimately facilitating instrument navigation and maneuvering should be a priority. Although advances in stereoscopic surgery aim to improve 3D perception, such systems have practical limitations with respect to their practical use as they tend to be extremely expensive and not widely available. For these reasons, it is useful to investigate other alternatives for conveying depth information and enhancing existing monocular visual cues. One of the primary cues that the visual system utilizes to infer depth is shadow [5]. Shadow can provide useful information about object shapes, relative 3D position, and surface characteristics within a scene [6,7,8,9]. Unfortunately, this visual cue is unavailable with MIS due to the coaxial alignment of the lens and light bundle of traditional rigid endoscopes. Under this setup, the operative field is generally shadowless [10]. It has been shown that inducing shadow by the introduction of a secondary light source within the surgical field improves endoscopic task performance [11]. The purpose of this paper is to introduce a new framework for improving depth perception for MIS by introducing a secondary light source. The unique feature of the proposed method is that it only casts a weak shadow of the laparoscopic instrument that is almost invisible under normal viewing conditions. During instrument maneuver, this "invisible shadow" is dynamically enhanced which introduces a strong depth cue from which

the distance between the instrument and tissue can be accurately determined. This naturally avoids the use of the instrument to "crash" on the tissue surface to gauge the 3D relative position of tissue-instrument whilst maintaining normal laparoscope viewing condition when instrument depth cuing is not required.

2   Materials and Methods

To emulate the laparoscopic environment, a laparoscopic box trainer was used, within which a silicon based tissue surface model was placed. The surface was coated with silicone rubber mixed with acrylic to give it a specular finish that looks similar to wet tissue. The scene was illuminated primarily from the endoscope itself and a secondary light source was placed directly above the surface but away from the endoscope. The intensity of this source was carefully adjusted so as to cast a near invisible shadow to the surface to avoid any interference to the normal viewing condition. A laparoscopic instrument was held over the surface in different positions and the vertical distance from the tip to the surface was measured to the nearest half centimeter. A video stream of the instrument assuming each position was obtained from a camera within the endoscope and digitally stored for subsequent processing. Figure 1 outlines the main components used for dynamically enhancing the "invisible shadow" during instrument maneuver.
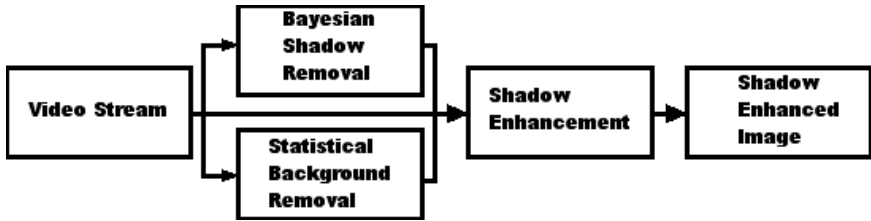

Fig. 2. A schematic illustration of the shadow enhancement filter design to generate the computer enhanced shadow

With this work, the shadow removal algorithm was based on the following four low level visual cues: intensity difference, intensity gain, angle between RGB vectors and color difference. Intensity difference is the absolute difference between the current image and the statistical background image B(x,y) calculated from the peak PDF of each pixel, $D(x,y)=|I(x,y)-B(x,y)|$ where $D(x,y)$ is the filter output. The use of intensity difference is biased towards the extraction of shadows in bright regions. For shadows in darker regions, however, one has to rely on the relative intensity attenuation between $I(x,y)$ and $B(x,y)$, given by $G(x,y)= | I(x,y)/B(x,y) |$. Based on the property of shadow color invariance, two color filters working in the RGB space have also been adopted:

$$R(x, y) = \frac{\langle \overline{I} \bullet \overline{B} \rangle}{\|\overline{I}\|\|\overline{B}\|}$$

where $\overline{I}$ and $\overline{B}$ are the RGB vectors of the background images. The final filter uses a color invariant model and addresses the limited color quantization steps:

$$c_1 = \arctan\left(\frac{R_i}{\max(G_i, B_i)}\right) \qquad b_1 = \arctan\left(\frac{R_b}{\max(G_b, B_b)}\right)$$

$$c_2 = \arctan\left(\frac{G_i}{\max(R_i, B_i)}\right) \qquad b_2 = \arctan\left(\frac{G_b}{\max(R_b, B_b)}\right)$$

$$c_3 = \arctan\left(\frac{B_i}{\max(R_i, G_i)}\right) \qquad b_3 = \arctan\left(\frac{B_b}{\max(R_b, G_b)}\right)$$

$$V(x, y) = (c_1 - b_1)^2 + (c_2 - b_2)^2 + (c_3 - b_3)^2$$

where (R_i, G_i, B_i) and (R_b, G_b, B_b) are the RGB components of a given pixel of the current background [12].

In order to evaluate the effect of shadow enhancement as an aid to depth perception, an experiment was devised comparing shadow enhanced to shadow unenhanced images. Ten volunteers with experiences in surgical imaging were recruited for this study. They were blinded to the aims of this study and asked to serially assess 36 images taken from the experimental setup described above on a 2D display. Each image showed part of laparoscopic instrument over a surface and the subjects were asked to estimate to the nearest half centimeter the vertical distance
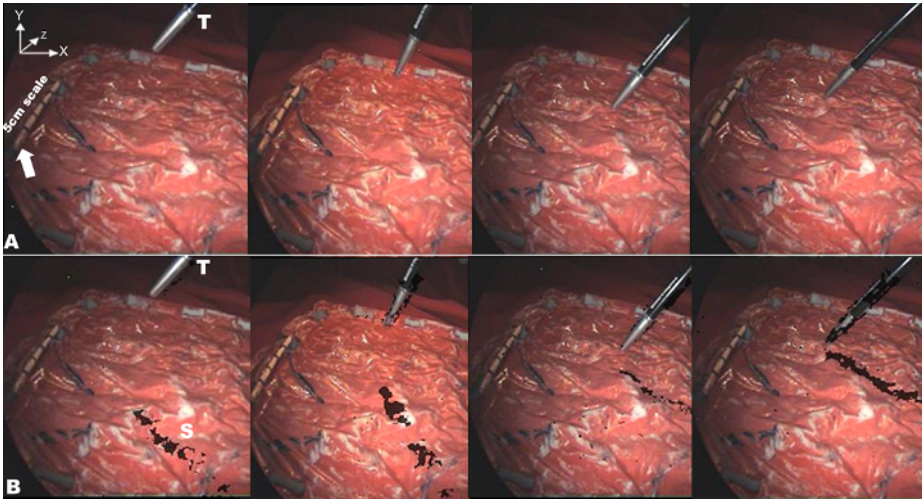

Fig. 3. (**A**) A series of raw images showing a laparoscopic instrument tip (T) as it approaches the silicon surface. (**B**) The same images following shadow (S) enhancement. Note the 5cm scaling aid to the left of each image (arrow).

from the tip of the instrument to the surface, with their answers recorded. The first 18 images were raw images (similar to those shown in Figure 3A) taken from the endoscope, whereas the remainder images were obtained by applying the described shadow enhancement algorithm (Figure 3B). For a better appreciation of the scaling and perspective, a 5cm marker with 1cm graduations was placed directly onto the surface in line with the z-axis.

To elucidate the underlying visual behavior of the users under normal and shadow enhanced viewing environments, gaze tracking was performed on all subjects performing the above task using a Tobii ET 1750 eye tracker. This is an infra-red video-based binocular eye-tracking system recording the position of gaze in the work plane (screen) at up to 38 samples per second with an accuracy of 1 degree across the work plane [13]. In this study the eye gaze data was analyzed qualitatively using the Clearview software (Tobii technology).

3 Results

To determine the perceptual accuracy, the absolute difference of the perceived distance of the tool-tip from the surface from the mean measured distance for all images for each subject was calculated. The mean difference for all subjects was 1.36cm for raw and 1.02cm for shadow enhanced images indicating an improvement in depth perception after shadow enhancement. This difference, however, was not statistically significant (t-test $p = 0.115$). The results indicate that the users are able to gauge large relative instrument distance to the surface when there are secondary visual cues. When the instrument is close to the tissue surface, however, the visual cues appreciable by the user are diminished if dynamic shadow enhancement is not applied. For this study, when the distance between the instrument and tissue surface is within 1cm (see Figure 4) (a situation that is most critical for relying on
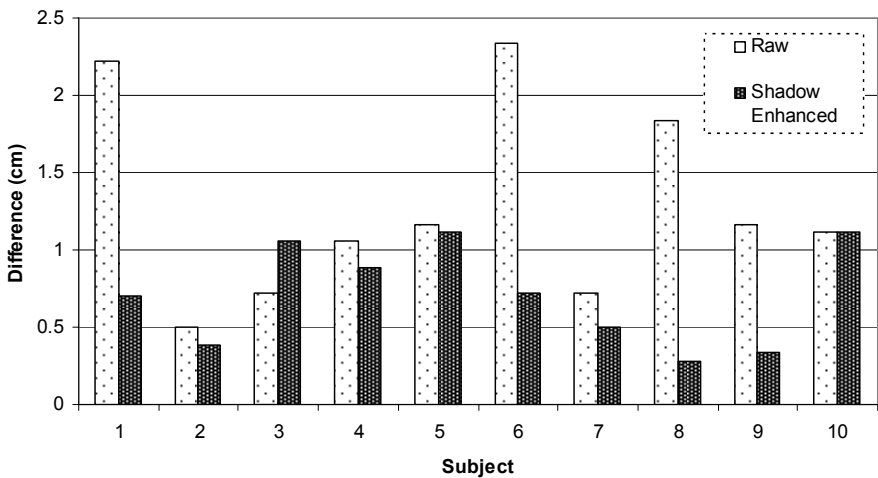

Fig. 4. Bar charts comparing the mean distance difference from reality perceived for raw and shadow enhanced images for each subject when the distance between the tool tip and tissue surface was set within 1cm
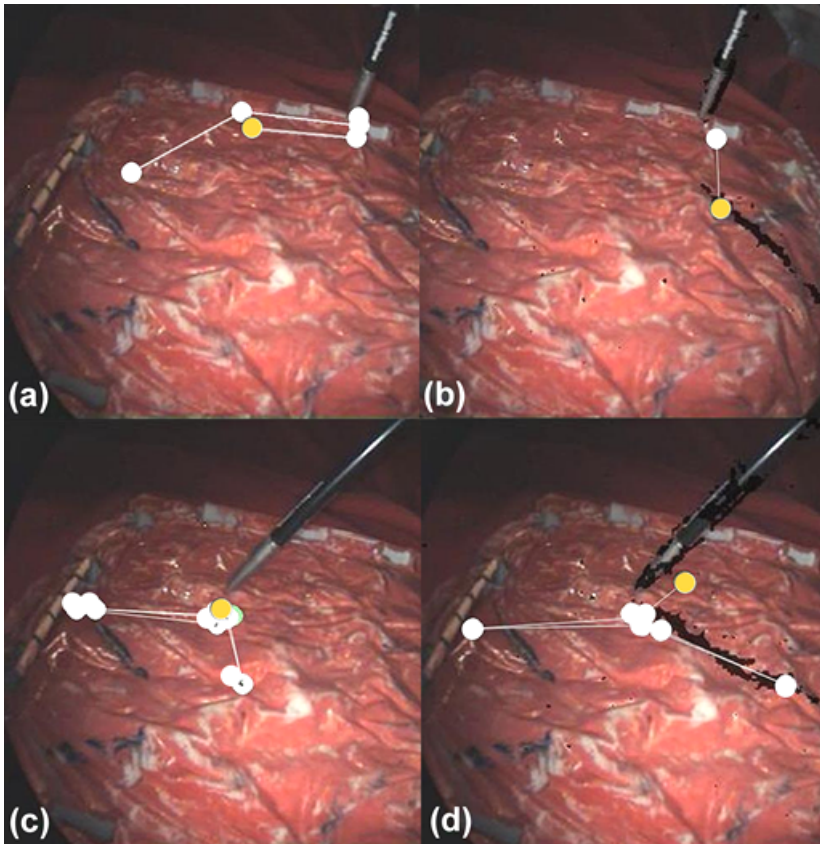
Fig. 5. An eye-gaze path demonstrating the fixations (white circles = fixations, yellow circle = first fixation, and white lines = saccades) recorded during the experiment. The effect of the "invisible shadow" enhancement to the accuracy of perceived tissue-instrument distance can be clearly demonstrated. In (a) and (b), the actual distance of the instrument tip from the tissue is 2cm and the subject whilst underestimating this to be 1cm in the raw image was able to estimate the exact distance correctly after shadow enhancement. In images (c) and (d) without shadow enhancement, the perceived distance was 2.5cm whereas by the use of shadow enhancement the perceived distance was 0.5cm, which is much closer to the ground truth.

tactile distance ranging), the mean errors for this group of observers were 1.283 and 0.709 (standard deviation 0.64 and 0.32) for raw and shadow enhanced images respectively (t-test p=0.020). This result was statistically significant. In addition, overall subjects were able to estimate depth faster from shadow enhanced images compared to raw images (5.8s versus 7.9s).

Verbal assessment of the participants has shown that all of them admitted that the enhanced "invisible" shadow had significantly facilitated their perception of depth. When a shadow was not present, most subjects based their answer on the scale of the instrument tip to estimate its position in space.

Figure 5 demonstrates the effect of the enhanced "invisible shadow" on general visual behavior revealed through eye tracking. It is evident that in Figures 5(a) and (b), the apex of the shadow provides direct cuing for depth perception with the subject drawing a visual line between the apex and tool tip. The real distance in this image is 2cm and the subject whilst underestimating to 1cm in the raw image was able to estimate the exact distance correctly after shadow enhancement. Figures 5(c) and (d), show a pair of images without and with "invisible shadow" enhancement when the instrument is in fact touching the surface. Without shadow enhancement, the perceived distance was 2.5 cm whereas by the use of shadow enhancement the perceived distance was 0.5cm, which is much closer to the ground truth.

4 Discussions and Conclusions

In this paper, we have demonstrated the effect of shadow on the accuracy of perceived tissue-instrument distance. One important feature of the algorithm is to cast an "invisible shadow" through the careful use of a secondary light source in a simulated laparoscopic environment. During instrument maneuver, this "invisible shadow" is dynamically enhanced which introduces a strong depth cue from which the distance between the instrument and tissue can be accurately determined. From a practical point of view, a faint shadow can theoretically be easily created by the introduction of a secondary light source through one of the accessory laparoscopic ports inserted during the procedure. The method naturally avoids the use of instrument to "crash" on the tissue surface, which is undesirable under delicate surgical maneuvers that require subtle control of the instruments to avoid damaging the tissues in contact. From both the objective and subjective assessment results of the study, it is evident that artificial shadow enhancement can be a useful aid for the perception of depth from 2D cues. Furthermore, the digital enhancement approach proved to be most effective when the instrument is in close proximity to a surface, which is the most critical time for enhanced instrument maneuver as *in vivo* mal-navigation at this level may lead to accidental injury to sensitive tissues.

References

1. Hubona GS, Shirah GW, Jennings DK. The effects of cast shadows and stereopsis on performing computer-generated spatial tasks *Systems, Man and Cybernetics, Part A, IEEE Transactions on* 2004; 34(4): 483-493.
2. Gibson JJ. Perception of the visual world. ed. Westport, Conn: Greenwood Press; 1974: pp. 235.
3. Kelsey C. Detection of visual information. In: The Perception of Visual Information, W. Hendee and P. Wells, Eds. New York: Springer-Verlag, 1993; 30-51.

4. Tendick F, Jennings R., Stark L, Tharp G. Sensing and Manipulation Problems in Endoscopic Surgery: Experiment, Analysis, and Observation. 1993; *Presence*, 2(1), 66-81.

5. Wanger L, Ferwanda J, Greenberg D. Perceiving spatial relationships in computer-generated images. *Computer Graphics and Applications, IEEE* 1992; 12(3): 44-58.

6. Huertas A, Nevatia R. Detecting Buildings in Aerial Images. *Comput. Vision Graphics and Image Processing.* 1988; 41: 131-152.

7. Irvin R B, NcKeown D.M. Methods for exploiting the relationship between building and their shadows in aerial imagery. *IEEE Tran. On System, Man and Cybernetics*, 1989; 19 (6): 1564-1575.

8. Liow Y, Pavlidis T. Use of shadows for extracting buildings in aerial images. *Computer Vision, Graphics and Image Processing*. 1990; 49: 242-277.

9. Thompson B W, Checky M T, Kaemmerer W F. Shadow stereo-locating object boundaries using shadows. *National Conf. on Art. Intel*. 1987; 761-766.

10. Mishra RK, Hanna GB, Brown SI and Cuschieri A. Optimum shadow-casting illumination for endoscopic task performance. *Arch Surg*, 2004; 139: 889-892.

11. Hanna GB, Cresswell AB, Cuschieri A. Shadow depth cues and endoscopic task performance. *Arch Surg* 2002; 137(10): 1166-1169.

12. Lo BPL and Yang GZ. Neuro-fuzzy shadow filter. *Lecture Notes In Computer Science:* Proceedings of the 7th European Conference on Computer Vision, 2002; 2352: 381-392.

13. Tobii technology. User Manual, 2003, http://www.tobii.se.